Research article

A SARS-CoV-2 host infection model network based on genomic human Transcription Factors (TFs) depletion

Massimiliano Chetta [a],[*], Alessandra Rosati [b], Liberato Marzullo [b], Marina Tarsitano [a], Nenad Bukvic [c]

[a] Ospedale Antonio Cardarelli, O.C. di Genetica Medica e di Laboratorio, A.O.R.N. Cardarelli, Medical Genetics Laboratory, Building Y, Naples, Italy
[b] Department of Medicine, Surgery and Dentistry "Schola Medica Salernitana", University of Salerno, Baronissi, 84081, Italy
[c] Azienda Ospedaliero Universitaria Consorziale Policlinico di Bari, Piazza Giulio Cesare 11, Bari, Italy

ARTICLE INFO

ABSTRACT

In December 2019 a new beta-coronavirus was isolated and characterized by sequencing samples from pneumonia patients in Wuhan, Hubei Province, China. Coronaviruses are positive-sense RNA viruses widely distributed among different animal species and humans in which they cause respiratory, enteric, liver and neurological symptomatology. Six species of coronavirus have been described (HCoV-229E, HCoV-OC43, HCoV-NL63 and HCoV-HKU1) that cause cold-like symptoms in immunocompetent or immunocompromised subjects and two strains of sometimes fatal zoonotic origin that cause severe acute respiratory syndrome (SARS-CoV and MERS-CoV). The SARS-CoV-2 strain is the emerging seventh member of the coronavirus family, which is actually determining a global emergency.

In silico analysis is a promising approach for understanding biological events in complex diseases and due to serious worldwide emergency and serious threat to global health, it is extremely important to use bioinformatics methods able to study an emerging pathogen like SARS-CoV-2.

Herein, we report on in silico comparative analysis between complete genome of SARS-CoV, MERS-CoV, HCoV-OC43 and SARS-CoV-2 strains, to identify the occurrence of specific conserved motifs on viral genomic sequences which should be able to bind and therefore induce a subtraction of host's Transcription Factors (TFs) which lead to a depletion, an effect comparable to haploinsufficiency (a genetic dominant condition in which a single copy of wild-type allele at a locus, in heterozygous combination with a variant allele, is insufficient to produce the correct quantity of transcript and, therefore, of protein, for a correct standard phenotypic expression).

In this competitive scenario, virus versus host, the proposed in silico protocol identified the TFs same as the distribution of TFBSs (Transcription Factor Binding Sites) on analyzed viral strains, potentially able to influence genes and pathways with biological functions confirming that this approach could brings useful insights regarding SARS-CoV-2. According to our results obtained by this in silico approach it is possible to hypothesize that TF-binding motifs could be of help in the explanation of the complex and heterogeneous clinical presentation in SARS-CoV-2 and subsequently predict possible interactions regarding metabolic pathways, and drug or target relationships.

1. Introduction

A novel coronavirus SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2), responsible for the recent pneumonia epidemic is a positive-sense RNA virus (ssRNA +), belonging to the Nidovirales order and Coronaviridae family, defined by a helical symmetry core and characteristic crown morphology [1]. Coronaviruses are genetically classified into four major genera: Alphacoronavirus, Betacoronavirus, correlated with mammal infections, and Gammacoronavirus, same as Deltacoronavirus that predominantly infects birds [2]. SARS-CoV-2 is a seventh strain of six different kinds of human CoVs previously identified: HCoV-NL63 and HCoV229E, that belong to the Alphacoronavirus genus; HCoV-OC43, HCoVHKU1, SARS-CoV (Severe Acute Respiratory Syndrome Coronavirus), and MERS-CoV (Middle East Respiratory syndrome Coronavirus), which belong to the Betacoronavirus group [3].

---

* Corresponding author.
  E-mail address: mchetta@unisa.it (M. Chetta).

The genome of SARS-CoV-2, is approximately 30,000 bases in length and is organized in 13–15 open reading frames (ORF) encoding a total of 27 proteins [2].

Like all coronaviruses the genome of SARS-CoV-2 encodes a complex of non-structural proteins polymeric RNA function (RNA-dependent RNA polymerase) and four main structural proteins: spike (S), membrane (M), envelope (E) and nucleocapsid (N) [2]. Additionally some other non-structural proteins are encoded in different numbers with uncertain functions. S glycoprotein, exposed on the virus surface, is the largest structural protein encoded by coronaviruses and is responsible for the binding and fusion of the virus with the host cell's membrane receptor. The protein S is encoded by an ORF with a highly variable nucleotide sequence that induces the main cell-mediated immune response. The host tropism of the coronaviruses is mainly determined by this protein which is composed of two portions: the N terminal end (S1) which performs a binding function to the receptor and the C-terminal end (S2) which performs the functions of fusion and entry of the virus into the cell [2].

The receptor binding domain (RBD) is a stretch of about 200 amino acids, located at the S1 end; it has the function of binding to the specific receptor on the surface of the host cell and represents the main immunostimulant epitope of the S protein [4]. Although, numerous variations have been identified in the sequence of SARS-CoV-2, it remains quite similar to SARS-CoV (about 79%) and MERS-CoV (about 50%) [2].

As expected, the greatest number of mutations take place in the region RBD and cause the recognition of different molecules on the surface of the host cell as a receptor. Human ACE2, expressed on the surface of the cells from respiratory and gastrointestinal tract, has been identified as the main cellular receptor for SARS-CoV-2 [5].

The growing evidence of the high adaptability of coronaviruses to new host animals in ecological niches (mammals, including humans, and birds) and the remarkable frequency of recombination, with a high mutation rate, required the use of tools that can provide quick information on the evolution of these viruses. Furthermore, the complex interplay between viruses and the host transcription machinery used by viruses for their own benefits, together with RNA viral replication, seems of fundamental importance.

New evidence confirmed that many positive and negative single-strand RNA viruses, whose primary replication site is cytoplasmic, are able to sequester nuclear proteins or TF to facilitate the replication process and alter the function of host cells. One of the mechanisms by which viruses can achieve this result is the disruption of the nucleus-cytoplasmic traffic and induce a spatial redistribution of proteins from the nucleus to the cytoplasm.

Several viruses (picornaviruses as well as VSV - vesicular stomatitis virus), use this kind of trafficking alteration to redistribute proteins to the cytoplasm, increase their interaction with the Internal Ribosome Entry Site (IRES) and facilitate translation of the virus polyprotein [6].

Recently was demonstrated that alphavirus transcripts are able to avidly bind cellular HuR protein through high-affinity binding sites conserved in their 3′-UTR. These RNA viruses, thanks to their ability to produce abundant amounts of viral RNA, can use this "sponge effect" as part of their strategy to modify cellular gene expression and interfere with cellular processes [7].

Among the emerging mechanisms of possible interactions between virus and host cell, it has been shown that many viruses (i.e. coronavirus SARS, influenza virus A, alpha and gamma herpesvirus) are able to increase the degradation of the host mRNA through interaction with specific viral proteins. These proteins determine endonucleolytic degradation of cytoplasmic mRNA by inducing widespread downregulation of host gene expression. This mechanism, known as "host shutdown", allows viruses to rapidly reduce the gene expression of the host cell andto attenuate the immune response same as to recover proteins needed for viral replication [8].

Furthermore, it must be considered that although human cells contain the same genome, they can differ in the expression of genes. It is known that gene expression is subject to multiple effects both spatial and temporal and does not depend exclusively on the information contained in the coding sequence of the DNA. It has recently been shown that viruses by subtracting parts of the machinery of host cells can induce epigenetic and epitranscriptomic modifications of gene expression.

On the bases of these observations, we propose the hypothesis that the virus infection could be capable of inducing a subtraction of host Transcriptional Factors (TFs) which lead to a depletion, an effect comparable to haploinsufficiency (a genetic dominant condition in which a single copy of wild-type allele at a locus, in heterozygous combination with a variant allele, is insufficient to produce the correct quantity of transcript and, therefore, of protein, for a correct standard phenotypic expression). That means that RNA viruses, could be able to used this ability as extra opportunity of their strategy to usurp cellular gene expression and interfere with target cellular processes [7, 9]. In this competitive scenario, virus versus host, the analysis pipeline was performed using different bioinformatics friendly tools (available online free of charge) on the complete genome of SARS-CoV, MERS-CoV, HCoV-OC43 and SARS-CoV-2 strains, to identify the occurrence of specific conserved motifs capable to bind human TFs and subsequently predict their possible interplay.

2. Materials and methods

The analysis pipeline was performed using different bioinformatics tools available online and consists of four main steps:

1) Analysis of complete strains of SARS-CoV, MERS-CoV, HCoV-OC43 and SARS-CoV-2 to discover conserved motifs on RNA sequences. The analysis on *.FASTA sequences was performed using MEME (Multiple EM For Motif Elicitation). By default MEME chooses the width and number of occurrences of each motif automatically in order to minimize the 'E-value' of the motif, the probability of finding an equally well-conserved pattern in random sequences. Only motif widths between 6 and 50 are considered [10, 11].

2) All the obtained motifs were used as query for Tomtom (http://meme-suite.org/doc/tomtom.html), another tool of MEME suite that compared the newly identified motifs against a database of known motifs (i.e., JASPAR). JASPAR CORE is a database that contained a curated and non-redundant set of open data access collections of experimentally discovered and proven TFs binding sites. Tomtom ranked the motifs in the database and produced an alignment for each significant match searching one or more query motifs against one or more databases of human target motifs (and their reverse complements when applicable). The report for each query was a list of target motifs, ranked by p-value and for each match an E-value and a q-value has been produced. The q-value is the minimal false discovery rate at which the observed similarity would be considered significant. Tomtom estimated q-values from all the match p-values using the Benjamini and Hochberg method. By default, significance was measured by q-value of the match. A list of Human TFs that contained the common conserved domain were obtained for all the motif's queries. All the human TFs were reported in Figure 1 [12,13].

3) Comparison between the all TFs obtained from the MEME output to select unique TFs for the SARS-CoV-2 strain.

4) Analysis on STITCH (Search Tool for Interactions of Chemicals) - http://stitch.embl.de/), a bioinformatics tool capable of providing integrated information about possible interactions regarding metabolic pathways, crystal structures, binding experiments, and drug or target relationships (Figure 1) [14].

3. Results and discussion

The original analytical workflow was developed to screen possible conserved motifs on SARS-CoV-2 able to subtract human genome TFs and subsequently modify the regulatory host scale networks. The result of the
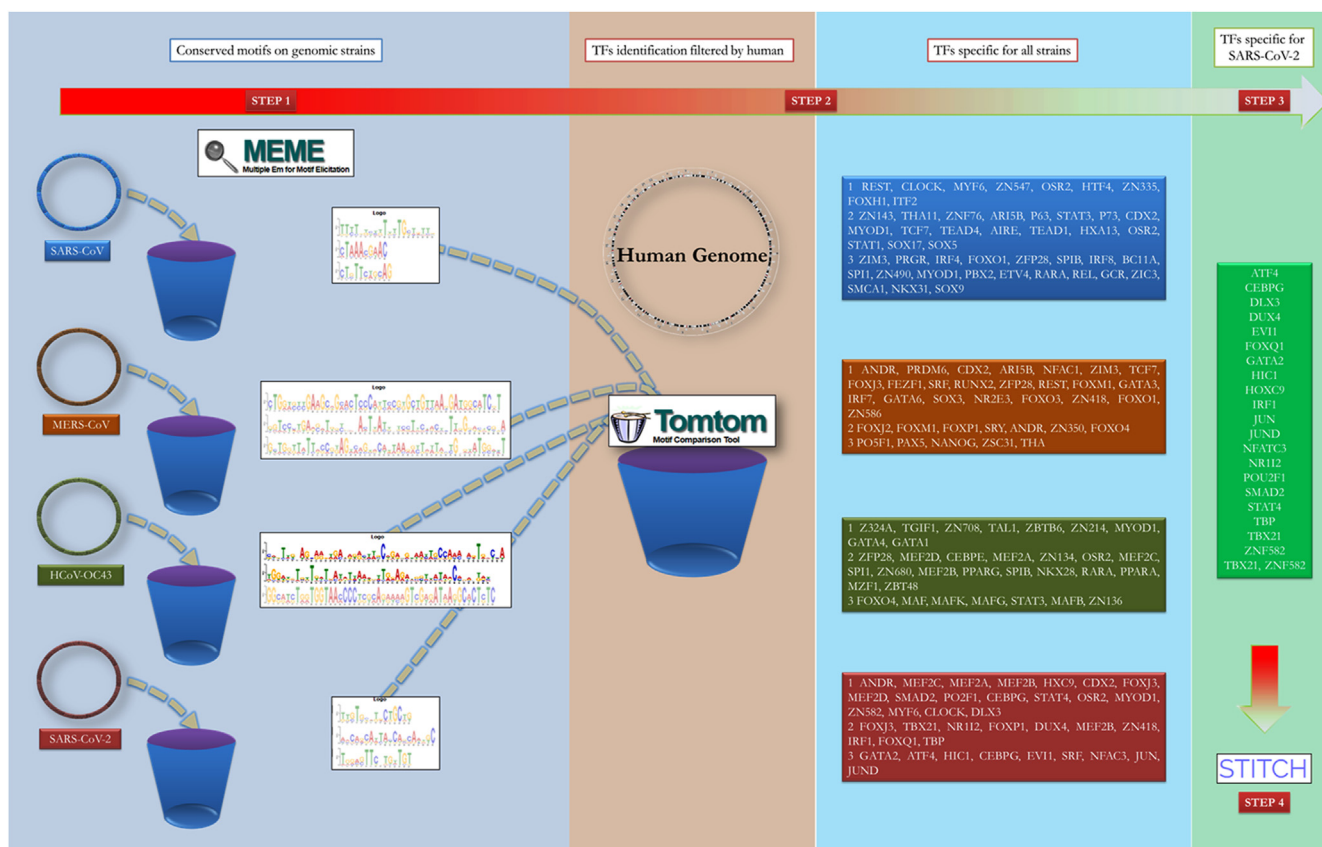
Fig. 1. A schematic representation of the pipeline integrating the different analytical software is provided. The details of the analytical steps are described in the text (materials and methods section).

in silico analysis permitted the identification of twenty unique TFs specific for SARS-CoV-2.

Using STITCH analysis, fifteen of those TFs were connected in a gene regulatory knowledge network, while five (HIC1 – HIC ZBTB Transcriptional Repressor 1, DUX4 - Double Homeobox 4, FOXQ1–Forkhead Box Q1, ZNF528 - Zinc Finger Protein 528 and NR1I2 - Nuclear Receptor Subfamily 1 Group I Member 2) seems to be without known interactions (Figure 2, panel A).

The approach adopted by our *in silico* analysis allowed to predict that a reduction of some TFs directly related to diffuse parenchymal lung disease, beside well known roles in the immune system. In particular, GATA2 (GATA-Binding Protein 2) deficiency is reported in children and adults with severe pulmonary alveolar proteinosis and hematologic disorders [15], while NFATC3 (Nuclear Factors of Activated T Cells Cytoplasmic, Calcineurin-Dependent 3) alteration is involved in proliferation of human pulmonary fibroblasts after hypoxic stimulus [16], and, finally, STAT4 (Signal Transducer and Activator of Transcription 4) associated to protective function in patients with systemic sclerosis [17]. The synergic action of these TFs could further clarify the mechanism of lung complications triggered by the infection of SARS-CoV-2.

Furthermore, the analytical workflow permitted to search indications for possible treatment on the basis of interaction between chemical compounds and TFs by selecting the function of chemical display directly on STITCH website. In particular, Tacrolimus, which is directly correlated with NFATC3 (Figure 2, panel B), has been identified. Tacrolimus used in therapy of patients for whom immunosuppression is needed, gives a promising results also in those patients with severe connective tissue interstitial lung disease [18]. With the same approach also Berberine a quaternary ammonium salt from the protoberberine group of benzylisoquinoline alkaloids has been individualized. This alkaloid reduces a fibroproliferative lung disorder in Idiopathic Pulmonary Fibrosis

(IPF) of unknown etiology by modulation of Transforming Growth Factor-β1 (TGF-β1)-mediated SMAD (SMAD Family Member 2; Mothers Against Decapentaplegic Drosophila Homolog 2) and non-SMAD signaling cascades (Figure 2 panel C) [19]. A very similar mechanism of action involving Janus Kinase 1 (JAK 1) and JAK2 has been described for the Ruxolitinib an oral inhibitor of JAK protein currently under investigation for possible use in treatment of COVID-19. These observations could be tested in biological systems and could be useful in processes towards the clarification of novel therapeutic aspects.

Moreover (Figure 2, panel A), our analysis indicates a close connection between different TFs capable of inducing a modulation of immune response. In particular, a deficiency of Activating Transcription Factor 4 (ATF4) that during Endoplasmic Reticulum (ER) stress-mediated inflammation, increases expression of inflammatory cytokines, including interleukin-6 (IL-6) and monocyte chemoattractant protein-1 (MCP-1), by inducing expression of NACHT, LRR, and PYD domain-containing protein 1 (NLRP1), a core component of the inflammasome [20]. The modification of interleukin (IL-6) expression, a multifunctional pro-inflammatory and anti-inflammatory cytokine, in turn [21], reduces promptly and transiently the response to infections and in particular to severe pneumonia induced by bacterial (i.e. S. Pneumoniae) [22]. The role of IL-6 as an anti-inflammatory cytokine is given, also, by its inhibitory effect on TNF-α (Tumor Necrosis Factor Alpha) and IL-1, same as by the activation of IL-10 and IL-1Ra (Interleukin 1 Receptor Antagonist) [23]. Furthermore, ATF4 is also implicated in the saturated fatty acid (SFA)-induced IL-6 expression in macrophages. Attenuation of ATF4 in macrophages markedly inhibits SFA-induced IL-6 expression [24]. Similarly is conceivable an inhibition of Interferon Regulatory Factor-1 (IRF1), the founding member of the interferon regulatory factor family, a transcription factor that regulates a range of target genes involved *in vivo* response to viruses and bacteria, and NFATC3 (Nuclear factor of
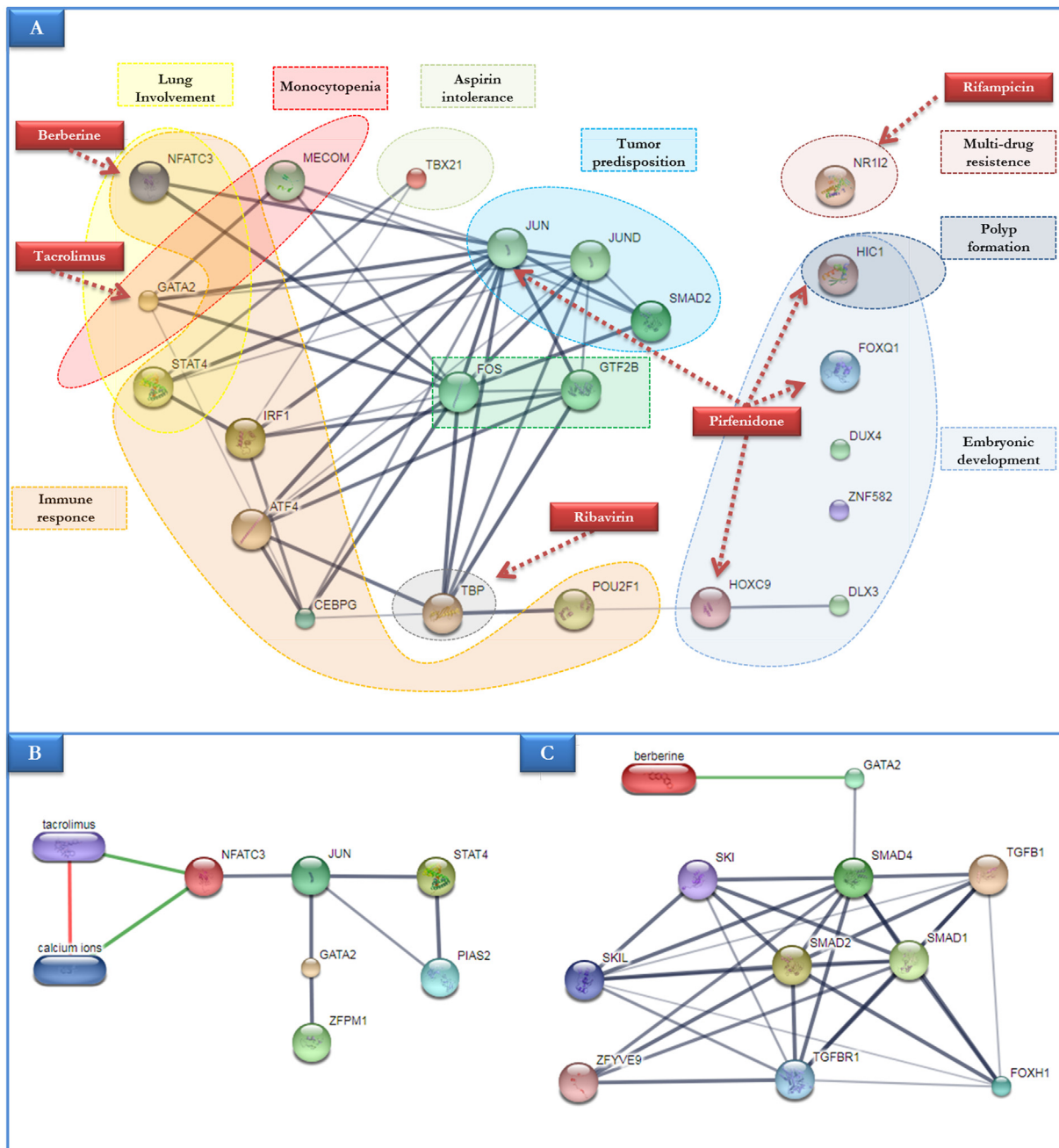
Fig. 2. The scheme provides an overview of the TFs computational relationship analysis (based on STITCH tool) associated with each functional clusters and drugs (panel A). Panel B and C shows the paths related to tacrolimus and berberine respectively.

activated T-cells, cytoplasmic 3) that plays a role in the induction of the IL-2 and expression of cytokine genes in T-cells [25, 26]. Through this approach CEBPG (CCAAT Enhancer Binding Protein Gamma) was identified too. This protein binds the promoter and the enhancer regions of target enhancer element PRE-I (Positive Regulatory Element-I) inducing the IL-4 gene expression, POU2F1 (POU Domain, Class 2, Transcription Factor 1) that modulates transcription of viral early genes immediately and STAT4 (Signal Transducer and Activator of Transcription 4) associated with a reduction of IFN-γ immunity and macrophage activity susceptibility [27, 28, 29].

It is particularly interesting that the modulation of the described TFs could be traced to the highest variability degree of IL1, IL6, IL10 and TNF-α related to the response against exogenous effects [30] and how these cytokine are part of three clusters with similar responses: first

[TNF-a, IL1, IL10], second [IFN-c, IL2, IL4, IL8, and IL12], and third [IL6]. Association of these cytokines and subsequent cellular adaptive immune system regulation and secretion are involved in the appearance of cytokine storm [30, 31, 32].

The analysis was expanded to characterize other possible implications related to the deficiency of other TFs (Figure 2, panel A). In particular, those TFs with apparent unknown interactions (HIC1, DUX4, FOXQ1, ZNF582), are involved in embryonic development, cell cycle regulation, tissue-specific gene expression, cell signaling, and tumorigenesis. Consequently, if infection occurs during precocious weeks of pregnancy a decrease of these TFs could be traced to a possible implication in embryonic development and possible fetal malformations [33, 34, 35, 36]. DLX3 (Distal-Less Homeobox 3) and HOXC9 (Homeobox C9), which, respectively, determine reduced placental formation of blood vessels and

increased pre and perinatal consequences [37, 38] could be added to these TFs. HIC1 is also correlated with allele accelerates polyp formation in adenomatous polyposis coli (APC) and in sporadic gastrointestinal (GI) cancers [39]. Of particular interest is NR1I2 (Nuclear Receptor Subfamily 1, Group I, Member 2), that when down-regulated, does not activate the transcription of multiple genes involved in the metabolism and secretion of potentially harmful xenobiotics (drugs and endogenous compounds). This down-regulation produces a multi drug resistance, specifically to rifampicin [40].

Interestingly, GATA2 (GATA Binding Protein 2) a transcription factor with an essential role in proliferation and differentiation of hematopoietic cells, as mentioned earlier, was identified among the TFs. An anomaly in GATA2 has been recently reported as the cause of familial syndromes with autosomal-dominant inheritance such as severe monocytopenia, NK and B lymphopenia, and absence of dendritic cells. Gata2 reduction also confers a selective advantage to EVI1 (Ecotropic Viral Integration Site 1) expression that contributes to an acceleration of leukemogenesis [41, 42].

Another group of TFs seems to be involved in possible development of human malignancies (see Figure 2, panel A,B,C). In particular the translocations and deletions of JUN (AP-1 transcription factor) and JUND (JunDProto-oncogene, AP-1 Tanscription Factor Subunit) related to deficiency of expression, is correlated to a development of predisposition of myelodysplastic syndrome and chronic myelomonocytic leukemia [43]. Besides these SMAD2 deficiency (SMAD Family Member 2) is related to an alteration of Tgfbr1 signaling and induced development of colorectal cancer [44].

Finally, TBP (TATA-Box Binding Protein) a member of evolutionarily conserved proteins known as TAF (TBP Associated Factors), a co-activator that facilitates RNA polymerase II complex assembly at the beginning of transcription, could be considered as another possible target for the development of therapy. TBP, is a target of Ribavirin a nucleoside analogue that interferes with viral synthesis of guanosine triphosphate. Although treatment with Ribavirin has been demonstrated to have no real efficiency on clinical outcome of patients with SARS or MERS, the possible direct interaction on TBP could contribute to reconsideration as a useful drug when administered alone or in combination with other compounds for COVID-19 [45,46].

In regards to the possibilities of selective therapeutic treatments of COVID-19 (SARS-CoV-2 infection), it seems possible that expression of four TFs (FOXQ1, HIC1, HOXC9, JUN), selected by our analysis, could be modulated by treatment with Pirfenidone (Figure 2, panel A). This drug reduces fibroblast proliferation, production of fibrosis-associated proteins and cytokines, as well as an increased biosynthesis and accumulation of extracellular matrix in response to cytokine growth factors such as Transforming Growth Factor - Beta (TGF-β) and Platelet - Derived Growth Factor (PDGF); used to treat idiopathic pulmonary fibrosis [47].

Finally, our analysis indicates that a possible deficiency of TBX21 (T-Box Transcription Factor 21) is associated with significant risk of Aspirin-induced asthma [48]. Therefore, at least in the early stages of the illness, it may be prudent to avoid use of this non-steroidal anti-inflammatory drug (NSAID).

4. Conclusion

In conclusion, the extended knowledge of TFs identified by this *in silico* approach with apparent deep translational relevance could provide an insight of the molecular aspects as they relate to the infection process, the dynamics of changes in the host's cellular system; clarifies the possible mechanism of elevated pulmonary tropism and, finally, indicate some clarifications that reflect on paving the way to the best therapeutic strategies. Even though further *in vitro* and *in vivo* analysis will be necessary to support the functional evidences derived from above reported observations of our *in silico* research, it seems reasonable to advise that all patients cured from SARS-CoV-2 infection, to undergo long term follow up

for tumor susceptibility due to possible oncogenic potential of SARS-CoV-2. Same attention should be paid in regards to the potential pregnancy risks.

Genetic variation is the process, devoid of specific and predetermined purposes, which underlies the differentiation and, therefore, the evolution of viruses. The set of viral mutations that originate as a consequence of different molecular mechanisms, allow the virus adaptation to the environment that will act by selecting the most suitable characteristics [49, 50]. RNA viruses, with high predisposition to replication error mediated by RNA-polymerase or reverse transcriptase are under the high diversification process [51]. The introduction of errors in the virus genome, generally around leading or consensus sequences, determines an extremely heterogeneous population structure which, in many cases, follows the concept of "quasispecies" - the viral population within a host composed by a distribution of different mutants in dynamic balance [52, 53]. This population is subject to continuous processes of genetic variation, competition and adaptation, that favor subsequent selection. Surface proteins that bind to the cell receptor, even if the non-structural proteins and the regulatory regions of the viral genome can influence the tropism of virus for the host cell play the main role in this selection [54]. Considering the evidence as well as the public health emergencies such as COVID-19 (SARS-CoV-2), it is necessary to find new rapid analysis strategies, such as computational approaches which, in addition to providing indications on the mechanisms of action of the virus, can also provide evolutionary models and possible models of transmission following the disturbance of the host's cellular network.

Even though *in silico* analyses require in vitro and in vivo validations, undoubtedly could provide valid contribution to gain deeper understanding of COVID-19 (SARS-CoV-2) and could be of help in global effort to find out effective measures, designing a suitable vaccine and/or therapeutic strategy [55, 56, 57].

Declarations

Author contribution statement

References

[1] A.R. Fehr, S. Perlman, Coronaviruses: an overview of their replication and pathogenesis, Methods Mol. Biol. 1282 (2015) 1–23.

[2] A. Wu, Y. Peng, B. Huang, X. Ding, X. Wang, P. Niu, J. Meng, Z. Zhu, Z. Zhang, J. Wang, J. Sheng, L. Quan, Z. Xia, W. Tan, G. Cheng, T. Jiang, Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China, Cell Host Microbe 27 (3) (2020 Mar 11) 325–328.

[3] K.G. Andersen, A. Rambaut, W.I. Lipkin, et al., The proximal origin of SARS-CoV-2, Nat. Med. (2020).

[4] R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, Y. Bi, X. Ma, F. Zhan, L. Wang, T. Hu, H. Zhou, Z. Hu, W. Zhou, L. Zhao, J. Chen, Y. Meng, J. Wang, Y. Lin, J. Yuan, Z. Xie, J. Ma, W.J. Liu, D. Wang, W. Xu, E.C. Holmes, G.F. Gao, G. Wu, W. Chen, W. Shi, W. Tan, Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding, Lancet 395 (10224) (2020 Feb 22) 565–574.

[5] H.P. Jia, D.C. Look, L. Shi, M. Hickey, L. Pewe, J. Netland, M. Farzan, C. Wohlford-Lenane, S. Perlman, P.B. McCray Jr., ACE2 receptor expression and severe acute respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia, J. Virol. 79 (23) (2005 Dec) 14614–14621. PubMed PMID: 16282461; PubMed Central PMCID: PMC1287568.

[6] J.A. Hiscox, The interaction of animal cytoplasmic RNA viruses with the nucleus to facilitate replication, Virus Res. 95 (1-2) (2003 Sep) 13–22.

[7] M.D. Barnhart, S.L. Moon, A.W. Emch, C.J. Wilusz, J. Wilusz, Changes in cellular mRNA stability, splicing, and polyadenylation through HuR protein sequestration by a cytoplasmic RNA virus, Cell Rep. 5 (4) (2013) 909–917.

[8] S. Gilbertson, J.D. Federspiel, E. Hartenian, I.M. Cristea, B. Glaunsinger, Changes in mRNA abundance drive shuttling of RNA binding proteins, linking cytoplasmic RNA degradation to transcription, Elife 7 (2018), e37663. Published 2018 Oct 3.

[9] K. Tsai, R.B. Cullen, Epigenetic and epitranscriptomic regulation of viral replication, Nat. Rev. Microbiol. (2020 Jun 12) 1–12.

[10] A. Harwig, R. Landick, B. Berkhout, The battle of RNA synthesis: virus versus host, Viruses 9 (10) (2017 Oct 21).

[11] T.L. Bailey, M. Boden, F.A. Buske, M. Frith, C.E. Grant, L. Clementi, J. Ren, W.W. Li, W.S. Noble, MEME SUITE: tools for motif discovery and searching, Nucleic Acids Res. 37 (Web Server issue) (2009 Jul) W202–W208.

[12] T.L. Bailey, DREME: motif discovery in transcription factor ChIP-seq data, Bioinformatics 27 (12) (2011 Jun 15) 1653–1659.

[13] S. Gupta, J.A. Stamatoyannopoulos, T.L. Bailey, W.S. Noble, Quantifying similarity between motifs, Genome Biol. 8 (2) (2007) R24. PubMed PMID: 17324271; PubMed Central PMCID: PMC1852410.

[14] M. Kuhn, C. von Mering, M. Campillos, L.J. Jensen, P. Bork, STITCH: interaction networks of chemicals and proteins, Nucleic Acids Res. 36 (Database issue) (2008 Jan) D684–D688. Epub 2007 Dec 15. PubMed PMID: 18084021; PubMed Central PMCID: PMC2238848.

[15] M. Griese, R. Zarbock, U. Costabel, J. Hildebrandt, D. Theegarten, M. Albert, A. Thiel, A. Schams, J. Lange, K. Krenke, T. Wesselak, C. Schön, M. Kappler, H. Blum, S. Krebs, A. Jung, C. Kröner, C. Klein, I. Campo, M. Luisetti, F. Bonella, GATA2 deficiency in children and adults with severe pulmonary alveolar proteinosis and hematologic disorders, BMC Pulm. Med. 15 (2015 Aug 12) 87.

[16] L.K. Senavirathna, C. Huang, X. Yang, M.C. Munteanu, R. Sathiaseelan, D. Xu, C.A. Henke, L. Liu, Hypoxia induces pulmonary fibroblast proliferation through NFAT signaling, Sci. Rep. 8 (1) (2018 Feb 9) 2709.

[17] L. Yi, J.C. Wang, X.J. Guo, Y.H. Gu, W.Z. Tu, G. Guo, L. Yang, R. Xiao, L. Yu, M.D. Mayes, S. Assassi, L. Jin, H.J. Zou, X.D. Zhou, STAT4 is a genetic risk factor for systemic sclerosis in a Chinese population, Int. J. Immunopathol. Pharmacol. 26 (2) (2013 Apr-Jun) 473–478. PubMed PMID: 23755762; PubMed Central PMCID: PMC4105920.

[18] S. Bremer, N.T. Vethe, M. Skauby, M. Kasbo, E.D. Johansson, K. Midtvedt, S. Bergan, NFAT-regulated cytokine gene expression during tacrolimus therapy early after renal transplantation, Br. J. Clin. Pharmacol. 83 (11) (2017 Nov) 2494–2502.

[19] Y. Hu, H. Fahmy, J.K. Zjawiony, G.E. Davies, Inhibitory effect and transcriptional impact of berberine and evodiamine on human white preadipocyte differentiation, Fitoterapia 81 (4) (2010 Jun) 259–268.

[20] H. Huang, G. Jing, J.J. Wang, N. Sheibani, S.X. Zhang, ATF4 is a novel regulator of MCP-1 in microvascular endothelial cells, J. Inflamm. (Lond) 12 (2015 Apr 17) 31.

[21] Y. Iwasaki, T. Suganami, R. Hachiya, I. Shirakawa, M. Kim-Saijo, M. Tanaka, M. Hamaguchi, T. Takai-Igarashi, M. Nakai, Y. Miyamoto, Y. Ogawa, Activating transcription factor 4 links metabolic stress to interleukin-6 expression in macrophages, Diabetes 63 (1) (2014 Jan) 152–161.

[22] M. Witzenrath, B. Gutbier, A.C. Hocke, B. Schmeck, S. Hippenstiel, K. Berger, T.J. Mitchell, J.R. de losToyos, S. Rosseau, N. Suttorp, H. Schütte, Role of pneumolysin for the development of acute lung injury in pneumococcal pneumonia, Crit. Care Med. 34 (7) (2006 Jul) 1947–1954. PubMed PMID: 16715037.

[23] J. Han, R.J. Kaufman, Physiological/pathological ramifications of transcription factors in the unfolded protein response, Genes Dev. 31 (14) (2017 Jul 15) 1417–1438.

[24] Y. Iwasaki, T. Suganami, R. Hachiya, I. Shirakawa, M. Kim-Saijo, M. Tanaka, M. Hamaguchi, T. Takai-Igarashi, M. Nakai, Y. Miyamoto, et al., Activating transcription factor 4 links metabolic stress to interleukin-6 expression in macrophages, Diabetes 63 (2014) 152–161.

[25] V. Narayan, M. Eckert, A. Zylicz, M. Zylicz, K.L. Ball, Cooperative regulation of the interferon regulatory factor-1 tumor suppressor protein by core components of the molecular chaperone machinery, J. Biol. Chem. 284 (38) (2009 Sep 18) 25889–25899.

[26] M. Vaeth, S. Feske, NFAT control of immune function: New Frontiers for an Abiding Trooper, F1000Res 7 (2018 Mar 2) 260.

[27] C. López-Rodríguez, L. Botella, CorbíAL, CCAAT-enhancer-binding proteins (C/EBP) regulate the tissue specific activity of the CD11c integrin gene promoter through functional interactions with Sp1 proteins, J. Biol. Chem. 272 (46) (1997 Nov 14) 29120–29126. PubMed PMID: 9360988.

[28] G.W. Ma, Y.K. Chu, H. Yang, X.H. Yan, E.G. Rong, H. Li, N. Wang, Functional analysis of sheep POU2F3 isoforms, Biochem. Genet. (2019 Dec 31).

[29] E. Kuroda, T. Kito, U. Yamashita, Reduced expression of STAT4 and IFN-gamma in macrophages from BALB/c mice, J. Immunol. 168 (11) (2002 Jun 1) 5477–5482. PubMed PMID: 12023341.

[30] L. Joshi, M. Ponnana, R. Sivangala, L.K. Chelluri, P. Nallari, S. Penmetsa, V. Valluri, S. Gaddam, Evaluation of TNF-α, IL-10 and IL-6 cytokine production and their correlation with genotype variants amongst tuberculosis patients and their household contacts, PloS One 10 (9) (2015 Sep 11), e0137727.

[31] H.H. Yiu, A.L. Graham, R.F. Stengel, Dynamics of a cytokine storm, PloS One 7 (10) (2012), e45027.

[32] P.L. Minciullo, A. Catalano, G. Mandraffino, M. Casciaro, A. Crucitti, G. Maltese, N. Morabito, A. Lasco, S. Gangemi, G. Basile, Inflammaging and anti-inflammaging: the role of cytokines in extreme longevity, Arch. Immunol. Ther. Exp. 64 (2) (2016 Apr) 111–126.

[33] S. Kany, J.T. Vollrath, B. Relja, Cytokines in inflammatory disease, Int. J. Mol. Sci. 20 (23) (2019 Nov 28) pii: E6008.

[34] H.P. Mohammad, W. Zhang, H.S. Prevas, B.R. Leadem, M. Zhang, J.G. Herman, C.M. Hooker, D.N. Watkins, B. Karim, D.L. Huso, S.B. Baylin, Loss of a single Hic1 allele accelerates polyp formation in Apc(Δ716) mice, Oncogene 30 (23) (2011 Jun 9) 2659–2669.

[35] A. Dandapat, L.M. Hartweck, D. Bosnakovski, M. Kyba, Expression of the human FSHD-linked DUX4 gene induces neurogenesis during differentiation of murine embryonic stem cells, Stem Cell. Dev. 22 (17) (2013 Sep 1) 2440–2448.

[36] A. Bieller, B. Pasche, S. Frank, B. Gläser, J. Kunz, K. Witt, B. Zoll, Isolation and characterization of the human forkhead gene FOXQ1, DNA Cell Biol. 20 (9) (2001 Sep) 555–561. PubMed PMID: 1747606.

[37] A.E. Armstrong, S. Gadd, V. Huff, D.S. Gerhard, J.S. Dome, E.J. Perlman, A unique subset of low- risk Wilms tumors is characterized by loss of function of TRIM28 (KAP1), a gene critical in early renal development: a Children's Oncology Group study, PloS One 13 (12) (2018 Dec 13), e0208936.

[38] A. Chui, N.A. Pathirage, B. Johnson, M. Cocquebert, T. Fournier, D. Evain-Brion, B. Roald, U. Manuelpillai, S.P. Brennecke, B. Kalionis, P. Murthi, Homeobox gene distal-less 3 is expressed in proliferating and differentiating cells of the human placenta, Placenta 31 (8) (2010 Aug) 691–697.

[39] C. Sauvegarde, D. Paul, L. Bridoux, A. Jouneau, S. Degrelle, I. Hue, R. Rezsohazy, I. Donnay, Dynamic pattern of HOXB9 protein localization during Oocyte maturation and early embryonic development in mammals, PloS One 11 (10) (2016 Oct 31), e0165898.

[40] J. Zhang, P. Kuehl, E.D. Green, J.W. Touchman, P.B. Watkins, A. Daly, S.D. Hall, P. Maurel, M. Relling, C. Brimer, K. Yasuda, S.A. Wrighton, M. Hancock, R.B. Kim, S. Strom, K. Thummel, C.G. Russell, J.R. Hudson Jr., E.G. Schuetz, M.S. Boguski, The human pregnane X receptor: genomic structure and identification and functional characterization of natural allelic variants, Pharmacogenetics 11 (7) (2001 Oct) 555–572. PubMed PMID: 11668216.

[41] C. Vicente, A. Conchillo, M.A. García-Sánchez, M.D. Odero, The role of the GATA2 transcription factor in normal and malignant hematopoiesis, Crit. Rev. Oncol. Hematol. 82 (1) (2012 Apr) 1–17.

[42] K. Kataoka, M. Kurokawa, Ecotropic viral integration site 1, stem cell self-renewal and leukemogenesis, Canc. Sci. 103 (8) (2012 Aug) 1371–1377.

[43] G. Trøen, V. Nygaard, T.K. Jenssen, I.M. Ikonomou, A. Tierens, E. Matutes, A. Gruszka-Westwood, D. Catovsky, O. Myklebost, G. Lauritzsen, E. Hovig, J. Delabie, Constitutive expression of the AP-1 transcription factors c-jun, junD, junB, and c-fos and the marginal zone B-cell transcription factor Notch2 in splenic marginal zone lymphoma, J. Mol. Diagn. 6 (4) (2004 Nov) 297–307. PubMed PMID: 15507668; PubMed Central PMCID: PMC1867488.

[44] R. Tjitro, L.A. Campbell, L. Basova, J. Johnson, J.A. Najera, A. Lindsey, M.C.G. Marcondes, Modeling the function of TATA Box binding protein in transcriptional changes induced by HIV-1 Tat in innate immune cells and the effect of methamphetamine exposure, Front. Immunol. 9 (2019 Feb 4) 3110.

[45] N.I. Fleming, R.N. Jorissen, D. Mouradov, M. Christie, A. Sakthianandeswaren, M. Palmieri, F. Day, S. Li, C. Tsui, L. Lipton, J. Desai, I.T. Jones, S. McLaughlin, R.L. Ward, N.J. Hawkins, A.R. Ruszkiewicz, J. Moore, H.J. Zhu, J.M. Mariadason, A.W. Burgess, D. Busam, Q. Zhao, R.L. Strausberg, P. Gibbs, O.M. Sieber, SMAD2, SMAD3 and SMAD4 mutations in colorectal cancer, Canc. Res. 73 (2) (2013 Jan 15) 725–735.

[46] G. Koren, S. King, S. Knowles, E. Phillips, Ribavirin in the treatment of SARS: a new trick for an old drug? CMAJ 168 (10) (2003 May 13) 1289–1292. PubMed PMID: 12743076; PubMed Central PMCID: PMC154189.

[47] C.N. Ravarani, G. Chalancon, M. Breker, N.S. de Groot, M.M. Babu, Affinity and competition for TBP are molecular determinants of gene expression noise, Nat. Commun. 7 (2016 Feb 2) 10417.

[48] G. Kwapiszewska, A. Gungl, J. Wilhelm, L.M. Marsh, H. ThekkekaraPuthenparampil, K. Sinn, M. Didiasova, W. Klepetko, D. Kosanovic, R.T. Schermuly, L. Wujak, B. Weiss, L. Schaefer, M. Schneider, M. Kreuter, A. Olschewski, W. Seeger, H. Olschewski, M. Wygrecka, Transcriptome profiling reveals the complexity of pirfenidone effects in idiopathic pulmonary fibrosis, Eur. Respir. J. 52 (5) (2018 Nov 22) pii: 1800564.

[49] M. Akahoshi, K. Obara, T. Hirota, A. Matsuda, K. Hasegawa, N. Takahashi, M. Shimizu, K. Nakashima, L. Cheng, S. Doi, H. Fujiwara, A. Miyatake, K. Fujita, N. Higashi, M. Taniguchi, T. Enomoto, X.Q. Mao, H. Nakashima, C.N. Adra, Y. Nakamura, M. Tamari, T. Shirakawa, Functional promoter polymorphism in the TBX21 gene associated with aspirin-induced asthma, Hum. Genet. 117 (1) (2005 Jun) 16–26. Epub 2005 Apr 2. PubMed PMID: 15806396.

[50] R. Sanjuán, P. Domingo-Calap, Mechanisms of viral mutation, Cell. Mol. Life Sci. 73 (23) (2016 Dec) 4433–4448. Epub 2016 Jul 8. Review. PubMed PMID: 27392606; PubMed Central PMCID: PMC5075021.

[51] S. Duffy, Why are RNA virus mutation rates so damn high? PLoS Biol. 16 (8) (2018 Aug 13), e3000003.

[52] S.L. Fishman, A.D. Branch, The quasispecies nature and biological implications of the hepatitis C virus, Infect. Genet. Evol. 9 (6) (2009 Dec) 1158–1167.

[53] E. Domingo, J. Sheldon, C. Perales, Viral quasispecies evolution, Microbiol. Mol. Biol. Rev. 76 (2) (2012 Jun) 159–216.

[54] M.S. Maginnis, Virus-receptor interactions: the key to cellular invasion, J. Mol. Biol. 430 (17) (2018 Aug 17) 2590–2611.

[55] E. Kim, G. Erdos, S. Huang, W.T. Kenniston, C.S. Balmert, Donahue Carey, V. Stalin Raj, W.M. Epperly, B.W. Klimstra, L.B. Haagmans, E. Korkmaz, L.D. Falo Jr., A. Gambotto, Microneedle array delivered recombinant coronavirus vaccines: immunogenicity and rapidtranslational development, EBioMedicine (2020).

[56] C. Chakraborty, A.R. Sharma, G. Sharma, M. Bhattacharya, S.S. Lee, SARS-CoV-2 causing pneumonia-associated respiratory disorder (COVID-19): diagnostic and proposed therapeutic options, Eur. Rev. Med. Pharmacol. Sci. 24 (7) (2020) 4016–4026.

[57] D.E. Gordon, et al., A SARS-CoV-2 protein interaction map reveals targets for drug repurposing, Nature (22 April 2020).